DNA Repair

Reading: Lehninger, pp. 949-958

Objectives:

1. To learn the types of mutations that can occur in DNA.

2. To learn classes of mutagens and how they act.

3. To understand the repair systems responsible for reversing DNA damage.

A cell has only one or two sets of genomic DNA, and unlike protein or RNA molecules which can be quickly replaced if damaged, the DNA molecules are irreplaceable. Maintaining the integrity of the DNA is therefore imperative and each cell has a dedicated set of DNA repair systems.

What are Mutations and Where do They Come From?

A mutation is a heritable change in the base sequence of DNA which can arise from errors during DNA synthesis, or post-replicatively as a result of spontaneous or mutagen-induced damage (Figure 1). Mutagens can be physical or chemical in nature, and can arise from artificial or natural causes.

What Types of Mutations are There?

1. Point Mutations - one base replaces another

a. Transitions: a purine replaces a purine, or a pyrimidine replaces a pyrimidine most chemical mutagens cause transitions

A-T T-A

C-G G-C

b. Transversions: a purine replaces a pyrimidine, or a pyrimidine replaces a purine; most transitions occur due to errors in replication

A-T T-A

C-G G-C

FIGURE 1
Mutation perpetuated by replication.
Mutations introduced on a DNA strand, such as the replacement of a cytosine by a uracil resulting from deamination of cytosine, extend to both strands when the damaged strand is used as a template during replication. In the first round of replication uracil selects adenine as complementary base. In the second round of replication uracil is replaced by thymine. Similar events occur when the other bases are altered.

2. Insertion/Deletion Mutations: One or more bps are inserted into or removed from the DNA. Large deletions or inversions can result in entire genes being duplicated or deleted

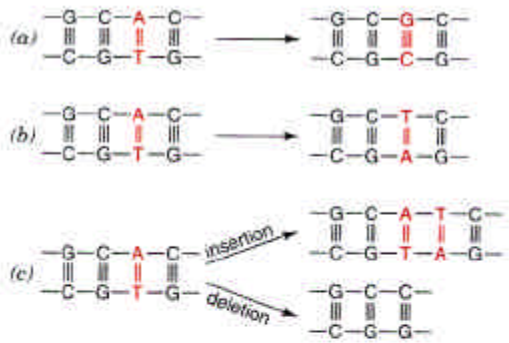


FIGURE 2
Mutations.
Mutations are classified as transition, transversion, and frameshift. Bases undergoing mutation are shown in color. (*a*) Transition: A purine–pyrimidine base pair is replaced by another. This mutation occurs spontaneously or can be induced chemically by such compounds as 5-bromouracil or nitrous acid. (*b*) Transversion: A purine–pyrimidine base pair is replaced by a pyrimidine–purine pair. This mutation occurs spontaneously and is common in humans. About one-half of the mutations in hemoglobin are of this type. (*c*) Frameshift: This mutation results from insertion or deletion of a base pair. Some insertions can be caused by mutagens such as acridines, proflavin, and ethidium bromide. Deletions are often caused by deaminating agents. Alteration of bases by these agents prevents pairing.

3. Inversions/Translocations: The DNA sequence has been flipped in orientation (inversion) or moved to an entirely new location in the genome (translocation).

What are the Consequences of Mutations?

Mutations in the DNA can lead to alterations in the function of the RNA and/or protein encoded; depending on the mutation this can lead to unaffected, partially impaired or completely destroyed activity of the encoded gene product. Point mutations occuring in protein coding sequences can cause amino acid changes as well as premature termination of the encoded protein while insertions and deletions can change the reading frame of the encoded protein (frameshift mutations). To fully appreciate the effects of various types of mutations on encoded proteins requires a complete understanding of how the information in DNA is converted to the amino acid sequence of proteins (discussed later). Mutations that confer no phenotype are "silent". Mutations in regulatory sequences can lead to altered levels of gene expression or to inappropriately-timed gene expression. A mutation

occurring in a germ cell will be passed to the offspring and, if harmful, will result in inheritable disease. Mutations that occur in the proliferating somatic cells are passed to all progeny cells.

Why are DNA Polymerases not More Accurate?

Many mutations arise from errors made by DNA Polymerase during DNA replication. DNA polymerases misincorporate bases at a frequency of 10^{-4} – 10^{-5}. The reason for this high rate of misincorporation is an inherent property of the bases themselves. The bases undergo keto to enol or amino to imino tautomerizations as a result of shifting of protons (Figure 3). In their rare tautomeric forms the bases have altered H-bonding properties that allow them to form non-Watson/Crick base pairs (Figure 4). In addition to the spontaneous mutations that result from misincorporation during replication, mutations also arise from spontaneous modifications of the DNA, most notably deamination of cytosine to uracil (Figure 5) and depurination.

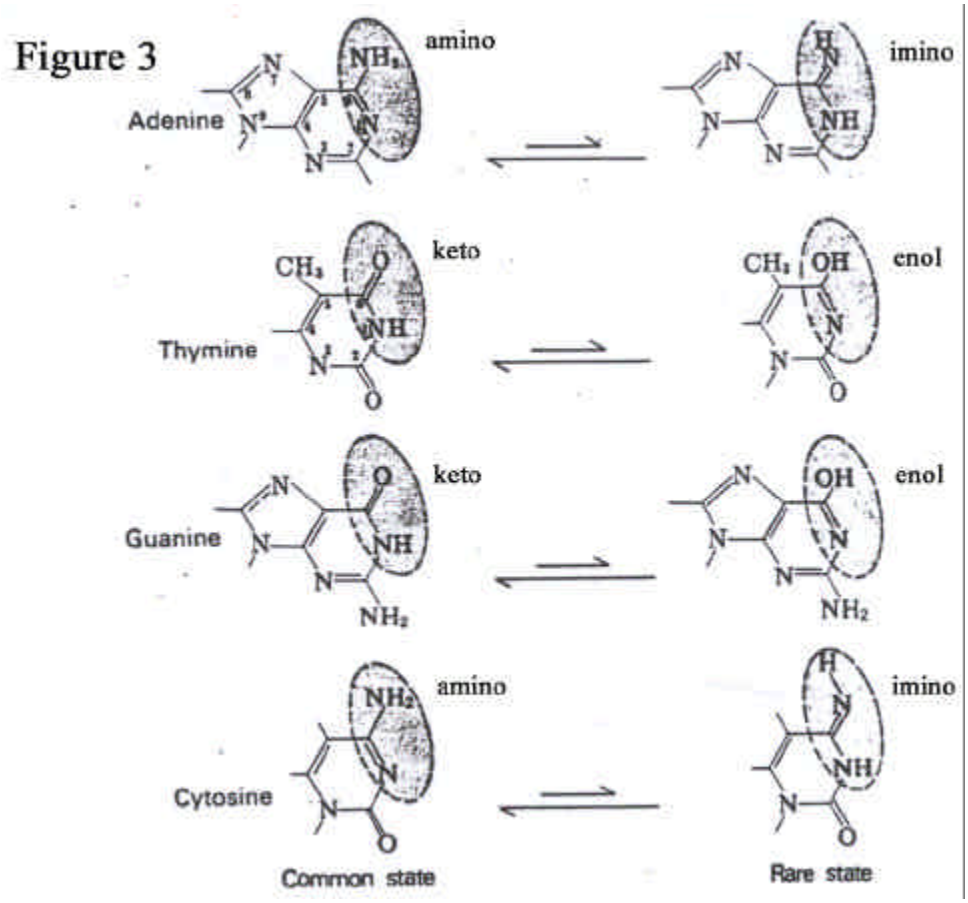


Figure 4

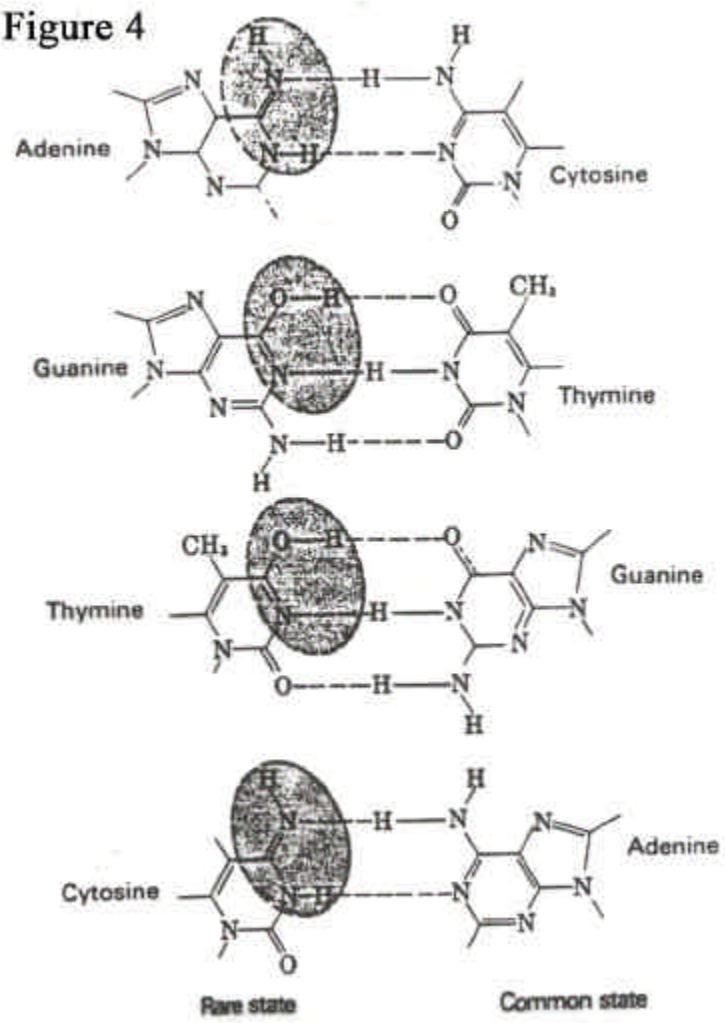
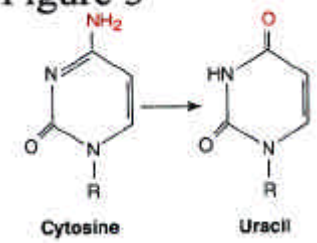


Figure 5



Cytosine spontaneously deaminates to Uracil at a high rate. Whereas cytosine should pair with guanine, uracil pairs with adenine (see Fig. 1).

How are Mutagens Identified?

The Ames test is a simple assay for mutagenic compounds. The test measures the potential of a given compound to promote easily detected mutations in a bacterial strain (Figure 6). Of compounds known to be carcinogenic from animal trials, more than 90% are also found to be mutagenic in the Ames test.

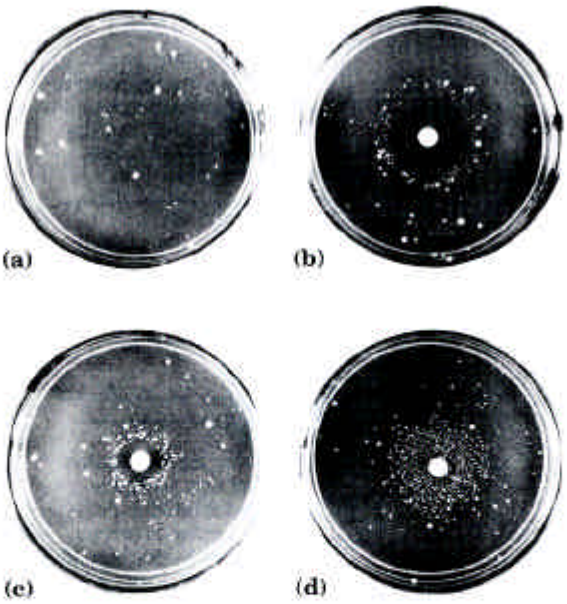


Figure 6

Ames test for carcinogens, based on their mutagenicity. A strain of *Salmonella typhimurium* having a mutation that inactivates an enzyme of the histidine biosynthetic pathway is plated on a histidine-free medium. Few cells grow. **(a)** The few small colonies of *S. typhimurium* that do grow on a histidine-free medium carry spontaneous back-mutations that permit the histidine biosynthetic pathway to operate. Three identical nutrient plates **(b)**, **(c)**, and **(d)** have been inoculated with an equal number of cells. Each plate then receives a disk of filter paper containing progressively lower concentrations of a mutagen. The mutagen greatly increases the rate of back-mutation and hence the number of colonies. The clear areas around the filter paper indicate where the concentration of mutagen is so high that it is lethal to the cells. As the mutagen diffuses away from the filter paper, it is diluted to sublethal concentrations that promote back-mutation. Mutagens can be compared on the basis of their effect on mutation rate. Because many compounds undergo a variety of chemical transformations after entering a cell, compounds are sometimes tested for mutagenicity after first incubating them with a liver extract. A number of compounds have been found to be mutagenic only after this treatment.

How do Mutagens Act?

Chemical Mutagens

1. Base Analogs

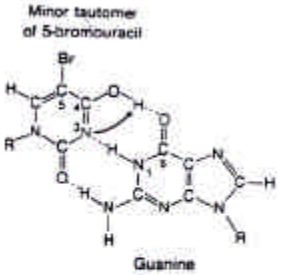


Figure 7

Hydrogen-bonding properties of the minor enol form of 5-bromouracil. The enol form of 5-bromouracil, an analog of thymine, pairs with guanine instead of adenine, the normal partner of thymine.

a. 2-aminopurine (2AP). This analog can pair with cytosine and thus cause A-T to G-C transitions.

b. 5-bromouracil (5BU) - 5BU can be incorporated into DNA in place of thymine. 5-BU has a higher tendency to exist in the enol form than thymine does, and in the enol form it base pairs with guanine (Figure 7) thereby causing A-T to G-C transitions.

2. Deaminating Agents - Nitrous acid (HNO_2) increases the rate of deamination of cytosine to uracil, guanine to xanthine and adenine to hypoxanthine (Figure 8). These changes lead to mutations because these modified bases form base pairs with partners other than the one the unmodified base would have. For example, although cytosine

would pair with guanine, deaminated cytosine (uracil) pairs with adenine (Figure 1). The common food preservative nitrite is the conjugate base of nitrous acid and thus forms the mutagen upon protonation. Nitrosamines are also potential mutagens because they react to generate nitrous acid.

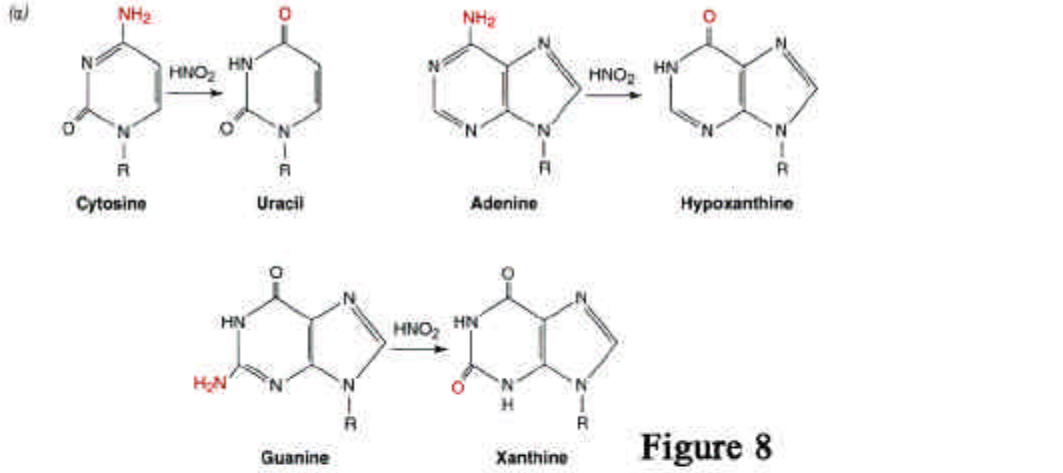
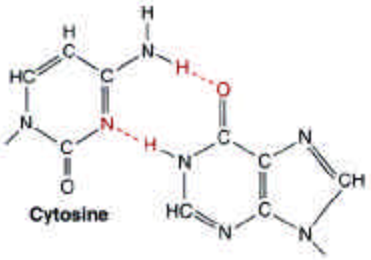


Figure 8

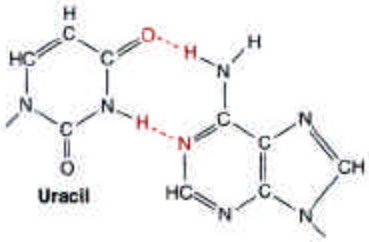
Reactions of various mutagens. (*a*) Deamination by nitrous acid (HNO_2) converts cytosine to uracil, adenine to hypoxanthine, and guanine to xanthine. (*b*) Reaction of bases with hydroxylamine (NH_2OH) as illustrated by the action of this reagent on cytosine. (*c*) Alkylations of guanine by dimethyl sulfate (DMS). Formation of a quaternary nitrogen destabilizes the deoxyriboside bond and releases deoxyribose. Among the effective agents for methylation of bases are nitrosoguanidines such as *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine.

3. Alkylating Reagents - include nitrogen mustard, ethylnitrosourea, dimethylsulfate and nitrosoguanidine. These agents cause transition mutations by altering the base pairing properties of purines. For example, ethylating N7 of guanine allows it to base pair with thymine resulting in a G-C to A-T transition (Figure 9). Furthermore, alkylation of purines renders the glycosidic bond more labile and promotes depurination. Since the apurinic site is a target for repair, mutations can be introduced as a result (see below).

Pairing of hypoxanthine with cytosine



Pairing of uracil with adenine



Pairing of 7-ethylguanine with thymine

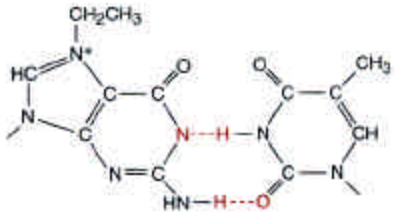


Figure 9

Chemical modifications that alter hydrogen-bonding properties of bases. Hypoxanthine, obtained by deamination of adenine, has different hydrogen-bonding properties from adenine and pairs with cytosine. Similarly, uracil obtained from cytosine has a different hydrogen-bonding specificity than cytosine and pairs with adenine. Alkylation of guanine modifies hydrogen-bonding properties of the base.

4. Intercalating Agents - include ethidium bromide and acridines. These are compounds that insert between the stacked bases of DNA causing distortion of the helix. Replication of the distorted regions can lead to base insertion or deletion (Figure 10).

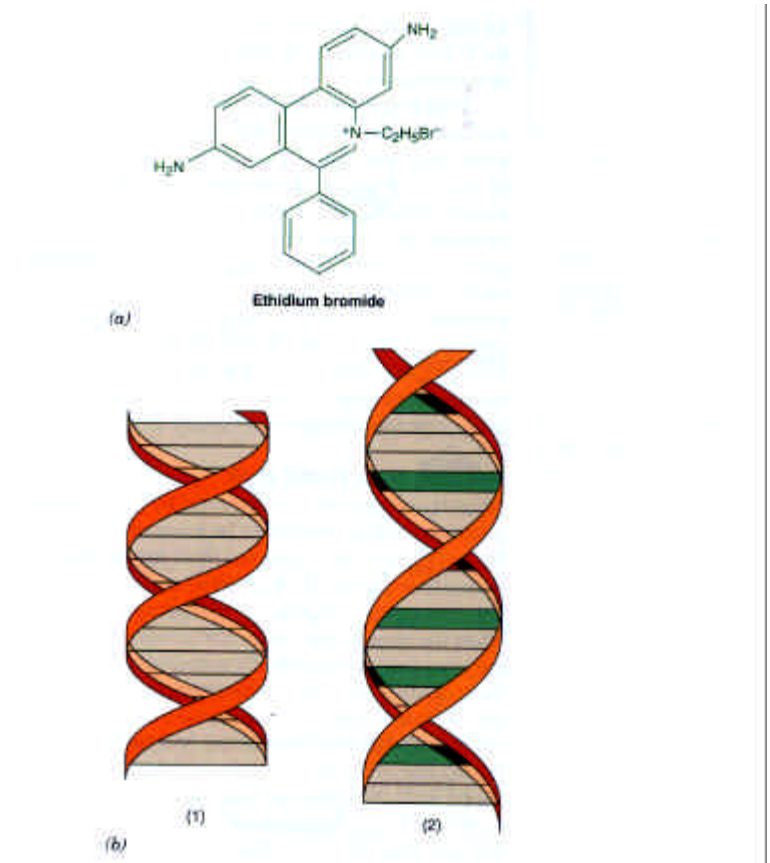


Figure 10

Intercalation between base pairs of the double helix. (*a*) Insertion of planar ring system of intercalators between two adjacent base pair requires stretching of the double helix (*b*). During replication this stretching apparently changes the frame used by DNA polymerase for reading the sequence of nucleotides. Consequently, newly synthesized DNA is frameshifted. (*b*-1) Original DNA helix; (*b*-2) helix with intercalative binding of ligands.

Redrawn based on figure in Lippard, S. J. Acct. Chem. Res. 11:211, 1978. Copyright © 1978 by the American Chemical Society.

Physical Mutagens

1. Ionizing Radiation - radiation can induce single- or double-stranded breaks. Radiation energy also can increase the proportion of the minor tautomers of the bases which can form non-Watson/Crick base pairs.

2. Ultraviolet Radiation - primarily affects the pyrimidines by activating the ethylene (C5-C6 bond) leading to dimerization of two adjacent pyrimidines (Figure 11). Thymines are particularly susceptible to formation of the cyclobutyl ring, but cytosine dimers and cytosine-thymine dimers also form. DNA polymerase cannot replicate through pyrimidine dimers.

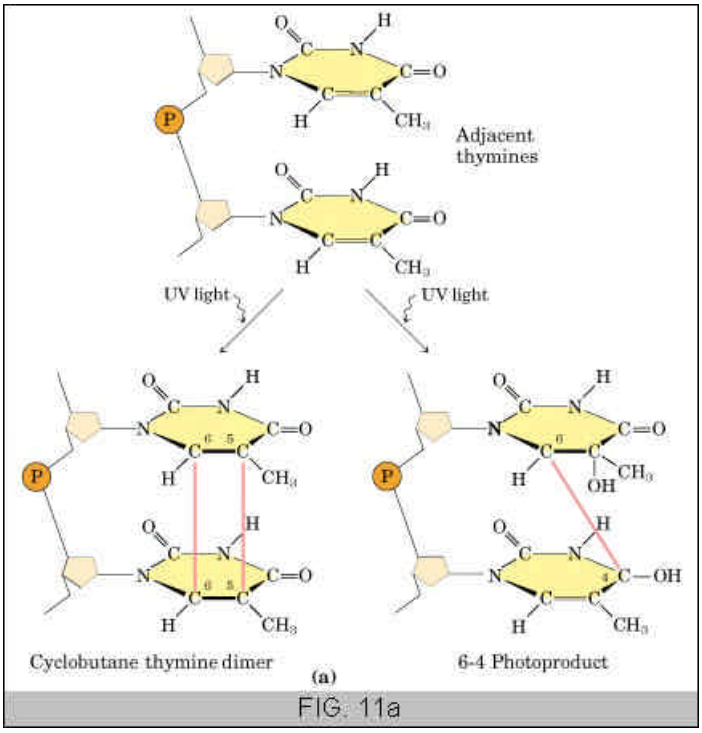


FIG. 11a

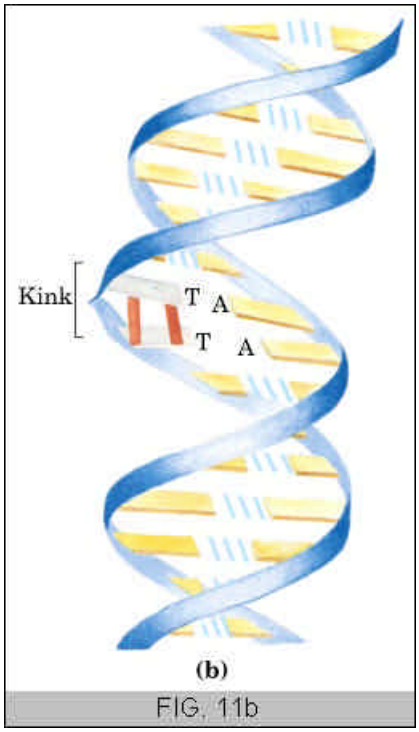


FIG. 11b

Table I DNA Lesions that Require Repair

DNA Lesion	*Cause*
Missing base	Acid and heat remove purines (~10^4 purines per day per cell in mammals)
Altered base	Ionizing radiation; alkylating agents
Incorrect base	Spontaneous deaminations: C → U, A → hypoxanthine
Deletion–insertion	Intercalating agents (e.g., acridine dyes)
Cyclobutyl dimer	UV irradiation
Strand breaks	Ionizing radiation; chemicals (bleomycin)
Cross-linking of strands	Psoralin derivatives (light-activated); mitomycin C (antibiotic)

Source: From Kornberg, A. *DNA Replication*. San Francisco: Freeman, 1980, p. 608.

Are There Mechanisms to Recognize and Repair DNA Damage?

The genome of a mammalian cell accumulates thousands of lesions a day. However, as a result of DNA repair, only 1/1000 becomes a mutation. A summary of lesions that can be repaired are presented in Table I.

Types of DNA Repair Systems in *E. coli*

Enzymes/proteins	Type of damage
Mismatch repair	
Dam methylase	Mismatches
MutH, MutL, MutS proteins	
DNA helicase II	
SSB	
DNA polymerase III	
Exonuclease I	
Exonuclease VII	
RecJ nuclease	
Exonuclease X	
DNA ligase	
Base-excision repair	
DNA glycosylases	Abnormal bases (uracil, hypoxanthine, xanthine); alkylated bases; pyrimidine dimers in some other organisms
AP endonucleases	
DNA polymerase I	
DNA ligase	
Nucleotide-excision repair	
ABC excinuclease	DNA lesions that cause large structural changes (e.g., pyrimidine dimers)
DNA polymerase I	
DNA ligase	
Direct repair	
DNA photolyases	Pyrimidine dimers
O^6-Methylguanine-DNA methyltransferase	O^6-Methylguanine

TABLE II

The number and diversity of repair systems (Table 2) reflect the importance of repair to cell survival. These repair systems are energetically expensive to the cell and this too indicates how critical the integrity of DNA is to the cell. For some types of lesions (e.g., pyrimidine dimers) several distinct repair systems can repair them. DNA repair is possible largely because DNA molecules consist of two strands containing redundant information - damaged DNA in one strand can be replaced using the template instructions in the undamaged strand.

How are Mismatched bps Recognized? How does the Cell Know Which Base in a Mismatch is Correct and Which is a Mistake?

The correction of mismatches that escape the 3' to 5' exonuclease proofreading of DNA Pol III improves the accuracy of the replication process 10^2 - 10^3-fold. To reduce the mutation frequency, these mismatches must be corrected to correspond to the information in the template strand. Thus this repair system must have a system to discriminate the parental and daughter strands. This is accomplished by "tagging" the old (template) DNA strand by methylation. The mismatch repair system of *E. coli* consists of several proteins (Table 2) that function in either strand discrimination or in repair.

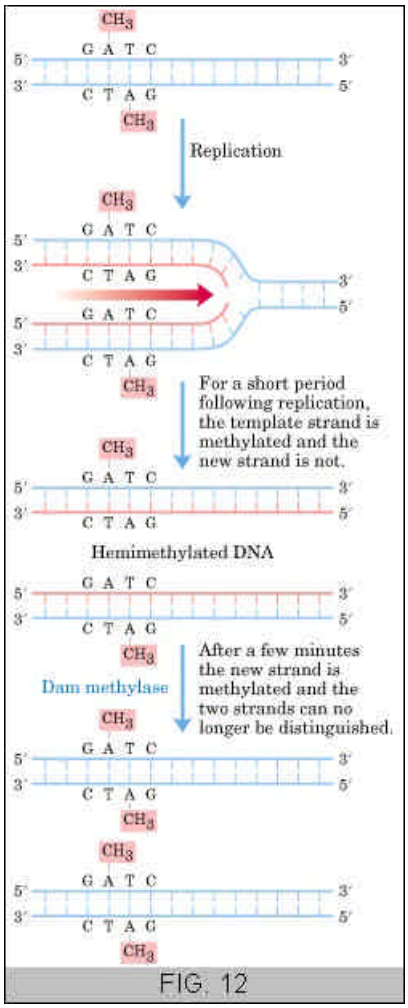


FIG. 12

The strand discrimating enzyme (DAM methylase) methylates DNA at the N6 position of adenines occurring in GATC sequences. For a short period after replication, the GATC sequences in the parental strand are methylated but the nascent strands are not (Figure 12). This provides the signal to specify which strand is parental and which is newly made. If both strands are methylated, no repair occurs. If neither strand is methylated, repair occurs but does not favor either strand. However when one strand is methylated and the other is not, enzymes recognize the mismatch, remove a stretch of DNA from the daughter (unmethylated) strand, and replace the DNA using the methylated strand as template. This repair system can correct mismatches up to 1000 bps from a partially (hemi)-methylated GATC. A model showing features of the repair is presented in Figure 13.

A complex of the MutS and MutL proteins binds to the mismatched base-pair (bp). MutH binds MutL and promotes threading of DNA on both sides of the mismatch through the MutS/MutL complex. This creates a loop and is equivalent to having the complex scan along the DNA on either side of the mismatch. MutH is a site specific endonuclease that cleaves hemimethylated GATC sequences on the unmethylated strand, marking the strand for repair.

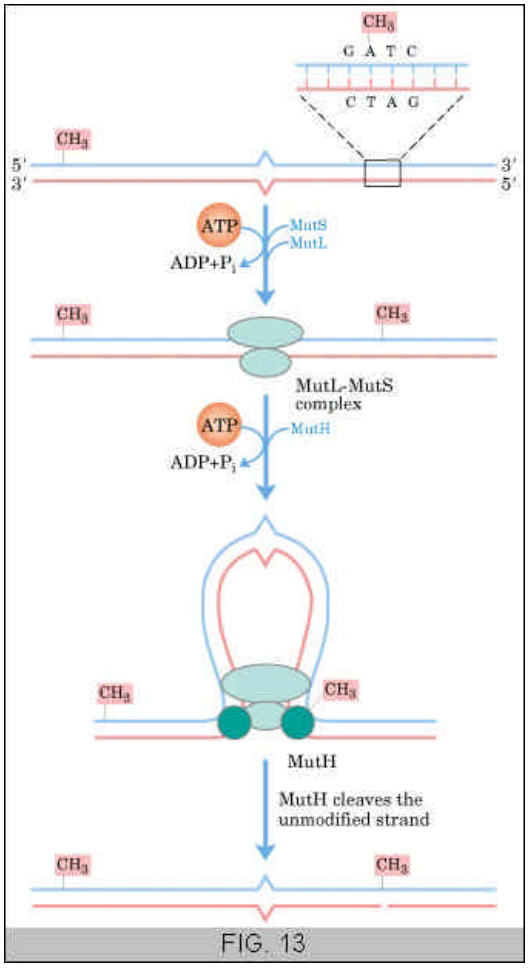


FIG. 13

The subsequent steps in the repair depend on whether the mismatch is on the 5' or 3' side of the cleavage site (Figure 14). If the mismatch is on the 5' side of the cleavage site (right panel of Fig. 14), DNA is degraded in the 3' to 5' direction past the mismatch by a 3'-exonuclease. If the mismatch is on the 3' side of the cleavage site (left panel of Fig. 14) the DNA is degraded in the 5' to 3' direction past the mismatch by a 5' exonuclease. In either case, once the mismatch is removed, the gap is repaired using DNA Pol III, and the nick is sealed using DNA ligase. A lot of energy is expended in degrading and resynthesizing 1000 bps of DNA, indicating the extreme importance of DNA integrity to the cell.

All eukaryotic cells have proteins that are analogous to the *E. coli* mismatch repair proteins, MutS and MutL. Mutations in the genes encoding these proteins produce some of the inherited cancer susceptibility syndromes. Mutations that compromise the DNA repair systems result in increased cancer frequencies.

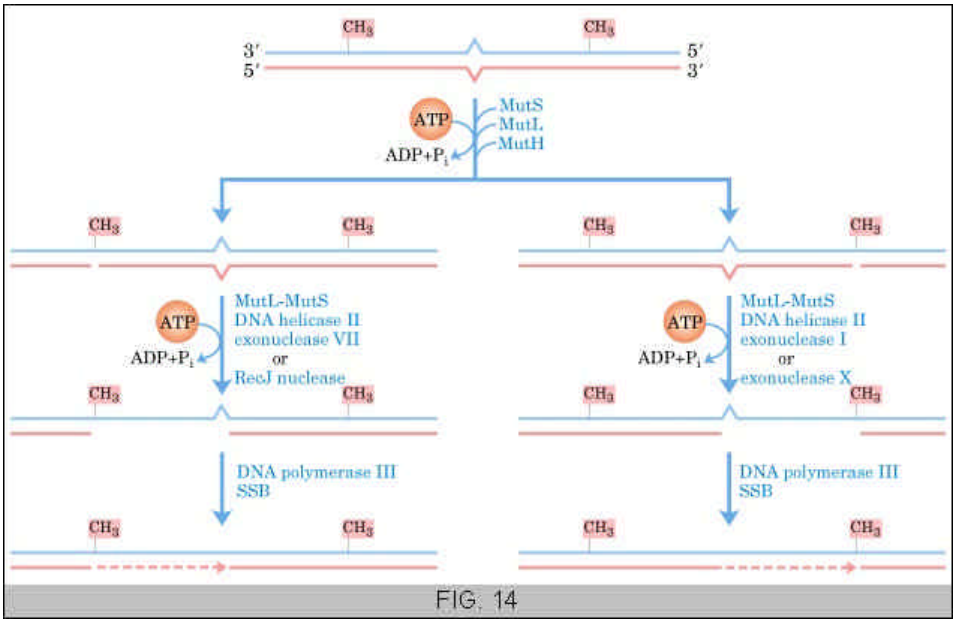


FIG. 14

Are other Lesions in the DNA (Chemically Altered Bases) Recognized?

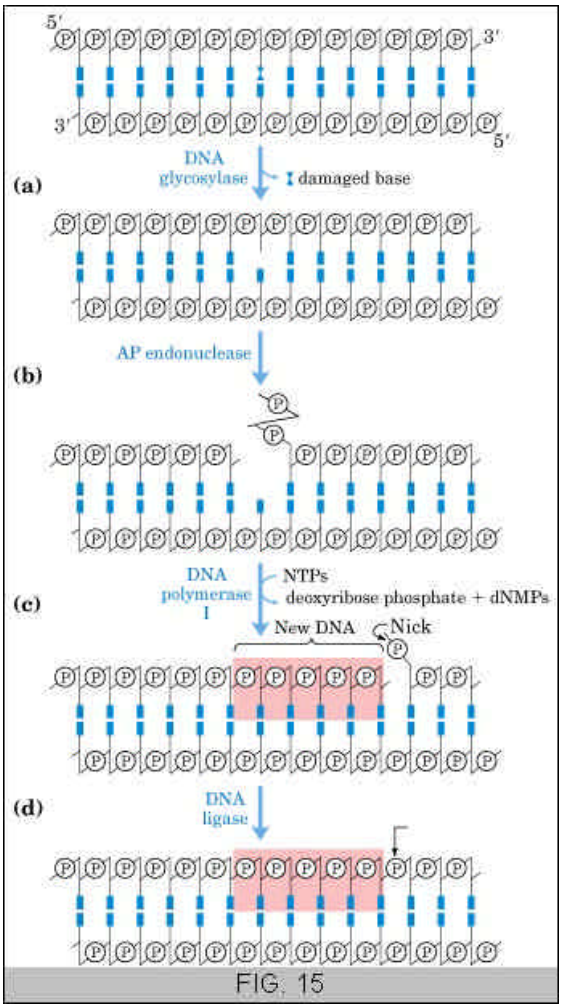


FIG. 15

Base Excision Repair

Another mechanism for recognizing and repairing DNA involves excision of the base and replacement (Figure 15). The DNA Glycosylases are a class of enzymes that recognize common lesions (eg., products of cytosine and adenine deamination) and remove the affected base by cleaving the N-glycosidic bond. This creates an apurinic or apyrimidinic (AP) site in the DNA. Each glycosylase is specific for a type of lesion. A common example is the uracil glycosylase that removes uracil arising from cytosine deamination. Other glycosylases remove hypoxanthine, alkylated bases, and pyrimidine dimers.

Once the AP site forms, another group of enzymes repairs it. Repair does not involve simply replacing the base. The strand containing the AP site is cut with an AP endonuclease, and a segment of DNA including the AP site is removed. The gap is repaired using DNA Pol I and DNA ligase. AP sites arising from slow spontaneous hydrolysis of the glycosidic bonds in DNA are also repaired by the AP endonuclease.

Nucleotide Excision Repair

DNA lesions causing distortions of the helix are repaired by Nucleotide-Excision Repair (Figure 16). In *E. coli* the key enzyme, ABC exinuclease, consists of three subunits (UvrA, UvrB and UvrC gene products). This enzyme recognizes many types of lesions including pyrimidine dimers and other adducts. A complex of UvrA/UvrB (A_2B) scans the DNA and when a distortion is encountered the complex binds and then UvrA dissociates. UvrC binds to UvrB and these two proteins cleave the phosphodiester chain on either side of the lesion. The protein encoded by uvrD has a helicase activity that unwinds and releases the excised oligontd. The gap is filled in with DNA Pol I and sealed by ligase. The mechanism of the mammalian nucleotide excision repair pathway is very similar (right panel of Fig 16), but in humans 16 proteins are required for the process. Genetic deficiencies in the nucleotide excision repair pathway in humans are responsible for serious diseases including Xeroderma Pigmentosum (to be discussed in the upcoming Clinical Correlation on DNA Repair Diseases).

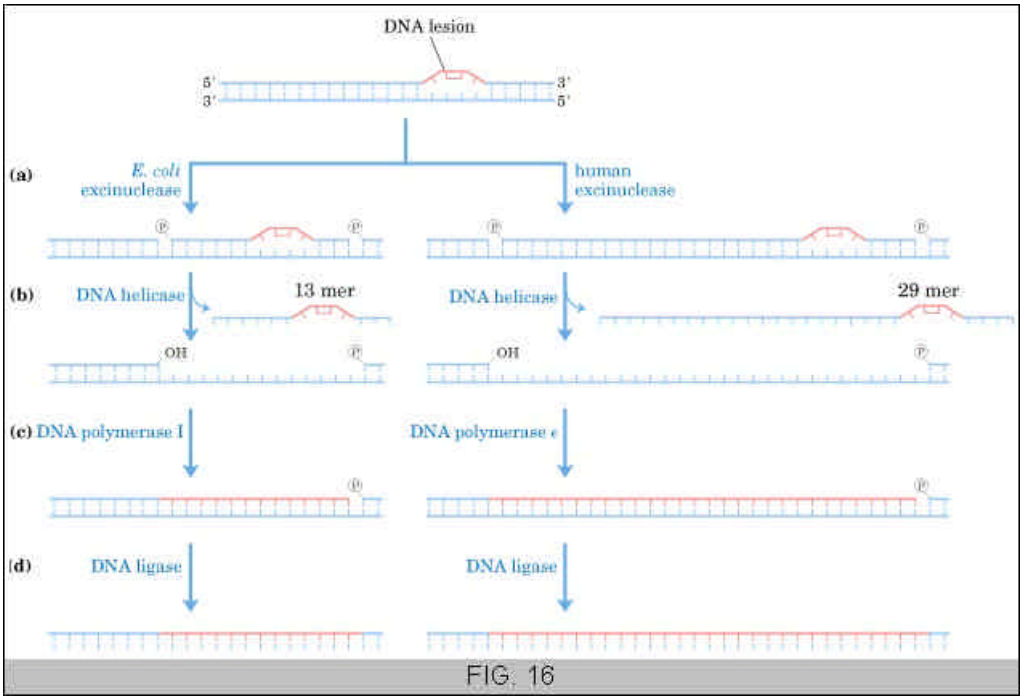


FIG. 16

Do any Repair Systems Simply Correct (rather than remove) the Lesion?

There are some examples of enzymes that act to directly repair damaged DNA.

1. Removal of pyrimidine dimers: thymine dimers can be removed by photoreactivating enzyme (also called photolyase). The enzyme binds directly to the dimer and cleaves the two bonds of the cyclobutyl ring to restore the normal pyrimidine structure. The reaction involves the absorption of visible light by the enzyme and its bound cofactors (Figure 17).

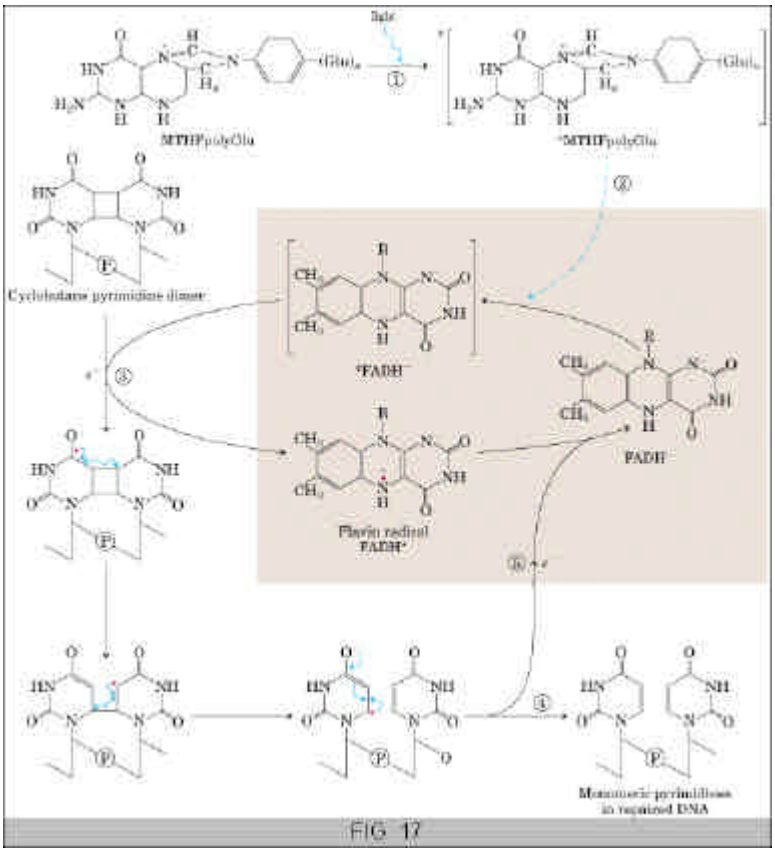
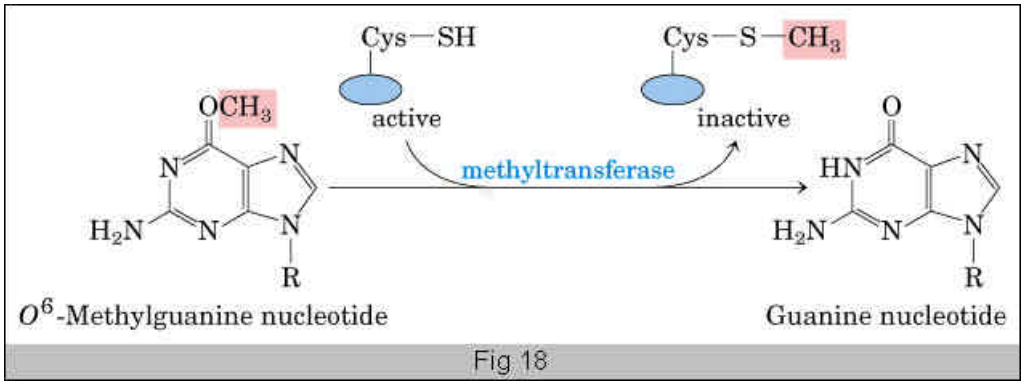


Figure 17

Repair of pyrimidine dimers with photolyase. Energy derived from absorbed light is used to reverse the photoreaction that caused the lesion. The two chromophores in *E. coli* photolyase (*M*, 54,000), N^5,N^{10}-methenyltetrahydrofolylpolyglutamate (MTHFpolyGlu) and $FADH^-$, perform complementary functions. On binding of photolyase to a pyrimidine dimer, repair proceeds as follows. ① A blue-light photon (300–500 nm wavelength) is absorbed by the N^5,N^{10}-methenyltetrahydrofolylpolyglutamate, which functions as a photoantenna. ② The excitation energy is transferred to $FADH^-$ in the active site of the enzyme. ③ The excited flavin ($^*FADH^-$) donates an electron to the pyrimidine dimer (shown here in a simplified representation) to generate an unstable dimer radical. ④ Electronic rearrangement restores the monomeric pyrimidines, and ⑤ the electron is transferred back to the flavin radical to regenerate $FADH^-$.

2. Dealkylation of alkylated bases: methylated and ethylated guanine residues can be repaired by direct transfer of the alkyl group from the O^6-methyl (ethyl) guanine to a specific cys residue on the O^6-Me guanine DNA methyltransferase (Figure 18). A single methyl transfer event permanently inactivates the protein. This is another example of the importance of repairing the DNA - an entire protein molecule is consumed in correcting one damaged base.

Fig. 18 Direct Repair by Removal of the Alkyl Group



Problems

1. What factors participate in ensuring the fidelity of replication during the synthesis of DNA?

2. How are potential mutagens tested?

3. Which repair systems can remove pyrimidine dimers from DNA?

4. When DNA Pol allows a mismatched base to survive in an elongating DNA molecule, how does the mismatch repair system know which base is correct and which is incorrect?

5. List some of the observations that lead to the conclusion that maintaining the integrity of the DNA is of utmost importance to the cell.

6. Compare and contrast the excision repair and photoreactivation repair mechanisms for correction of UV-induced thymine dimers.

7. Some sites in the *E. coli* chromosome are particularly susceptible to mutation. Many of these Ahot spots@ contain 5-methylcytosine residues. Suggest the reason for the high mutation rates at these sites.